

Research Article

Statistical Downscaling of ERA-Interim Forecast Precipitation Data in Complex Terrain Using LASSO Algorithm

Lu Gao,^{1,2,3} Karsten Schulz,⁴ and Matthias Bernhardt³

¹ College of Geographical Sciences, Fujian Normal University, Fuzhou 350007, China

² Fujian Provincial Engineering Research Center for Monitoring and Assessing Terrestrial Disasters, Fuzhou 350007, China

³ Department of Geography, Ludwig-Maximilians-Universität München, 80333 Munich, Germany

⁴ Institute of Water Management, Hydrology and Hydraulic Engineering, University of Natural Resources and Life Sciences, 1190 Vienna, Austria

Correspondence should be addressed to Lu Gao; l.gao@foxmail.com

Received 27 March 2014; Revised 21 June 2014; Accepted 23 June 2014; Published 10 July 2014

Academic Editor: Lian Xie

Copyright © 2014 Lu Gao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Precipitation is an essential input parameter for land surface models because it controls a large variety of environmental processes. However, the commonly sparse meteorological networks in complex terrains are unable to provide the information needed for many applications. Therefore, downscaling local precipitation is necessary. To this end, a new machine learning method, LASSO algorithm (least absolute shrinkage and selection operator), is used to address the disparity between ERA-Interim forecast precipitation data (0.25° grid) and point-scale meteorological observations. LASSO was tested and validated against other three downscaling methods, local intensity scaling (LOCI), quantile-mapping (QM), and stepwise regression (Stepwise) at 50 meteorological stations, located in the high mountainous region of the central Alps. The downscaling procedure is implemented in two steps. Firstly, the dry or wet days are classified and the precipitation amounts conditional on the occurrence of wet days are modeled subsequently. Compared to other three downscaling methods, LASSO shows the best performances in precipitation occurrence and precipitation amount prediction on average. Furthermore, LASSO could reduce the error for certain sites, where no improvement could be seen when LOCI and QM were used. This study proves that LASSO is a reasonable alternative to other statistical methods with respect to the downscaling of precipitation data.

1. Introduction

Precipitation is one of the most important variables for a large variety of environmental processes and its spatial and temporal variations directly influence the local as well as the global water, energy, and matter cycle (e.g., [1–6]). Precipitation is also an essential input parameter for land surface models in fields such as in hydrology, ecology, and climatology [7–9]. Traditionally, precipitation is measured by rain gauges. However, most of them are located in homogeneous terrains and in valley regions. In complex terrains, few gauge stations exist due to difficulties in snow depth measuring and maintenance of stations which result in a lack of long-term and high resolution records [10]. In order to overcome these limitations, in the past decades, GCMs have been widely applied to meet the specific needs

of environmental impact models by providing time series of precipitation and plausible scenarios of change [8, 11].

However, it is well known that the coarse spatial resolution (~300 km) of GCMs limits the reliable use of these data in decision making and model based impact studies [12–15]. Specifically, GCMs only provide precipitation characteristics that are based on the mean elevation of the pixel, thus not considering subgrid variability of topography and consequent atmospheric features. Local processes, such as orographic precipitation along mountain slopes, are not represented in the coarse grids [12]. Even if RCMs nested in GCMs provide better descriptions of local scale characteristics, based on the finer spatial resolution of 10–50 km, they do not fulfill the requirement of hydrological and climatic impact models, which typically run on the scale of 0.1–1 km [11, 16–21]. Thus, correction and downscaling are necessary

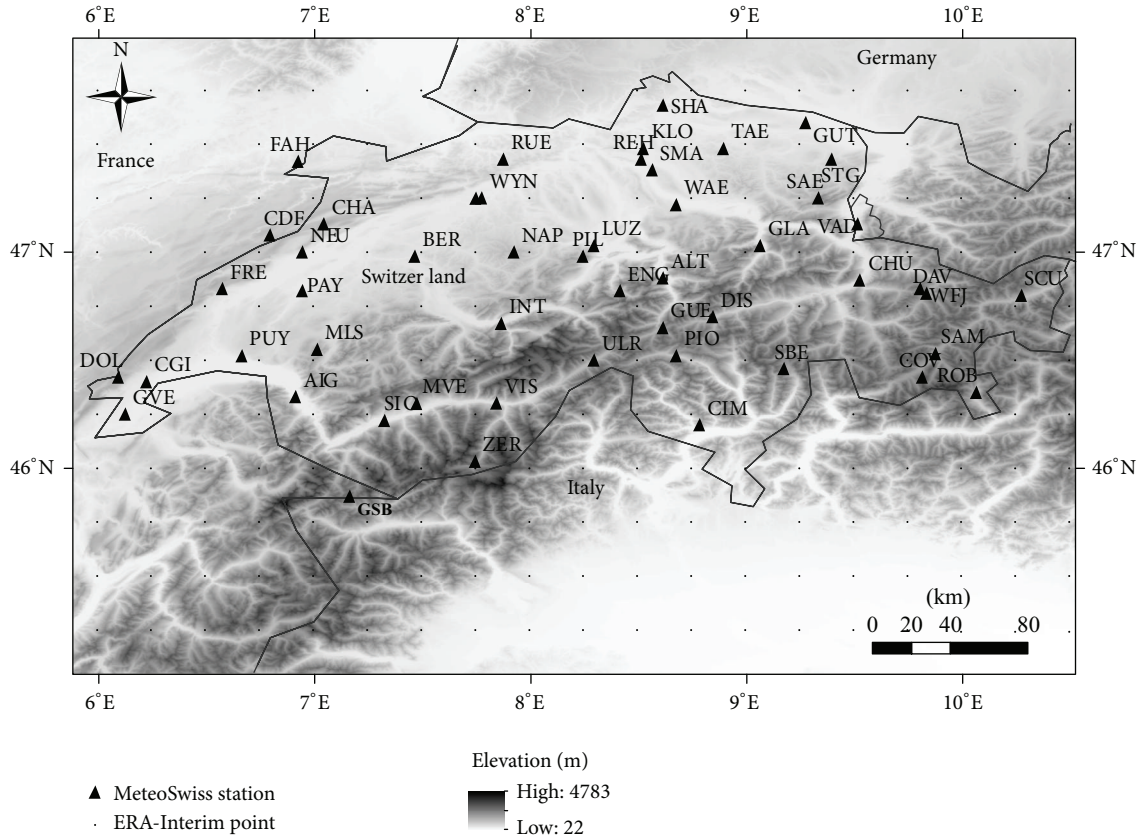


FIGURE 1: Location of the MeteoSwiss sites (triangles) and ERA-Interim $0.25^\circ \times 0.25^\circ$ points (dots). The elevation ranges from 22 to 4783 m a.s.l., with a DEM resolution of 90 m.

for both GCM and RCM outputs before they are applied for environment impact models [22–26].

In a broad sense, downscaling includes two purposes: one is obtaining future emission scenarios (time series) for stations or regions or catchments from global climate models (e.g., HadAM3P) using the established statistical relationship between measurements and large-scale predictors. The other purpose is obtaining time series for nonmeasurement area. This also could be treated as spatial interpolation or disaggregation to some extent. Many previous studies have shown that a reproduction of station data and the generation of future times series at the respective locations is extremely useful (e.g., [19, 22, 24]).

In terms of downscaling methods, dynamical and statistical downscaling are the two main approaches. The latter approach establishes the statistical connections between large-scale circulation variables (predictors) and local observed variables (predictands) [27, 28]. Compared to dynamical downscaling, statistical downscaling methods have lower computational demands and allow for a fast application [27, 29]. Maraun et al. [11] comprehensively reviewed precipitation downscaling methods from an end user's point-of-view. For example, local intensity scaling (LOCI) and quantile-mapping (QM) are the standard methods used for correcting GCM or RCM outputs with respect to local observations.

Other methods focus on investigating the relationship between local precipitation and large-scale atmosphere circulations, which vary from linear regression (e.g., [30–35]) to complex nonlinear models (e.g., [36–40]). Although numerous studies were carried out, a general standardized precipitation downscaling method still does not exist, especially for complex terrains. Furthermore, a separate predictor selection process (e.g., principal components analysis) is usually implemented in order to search for the most sensitive variables with regard to precipitation variations (e.g., [34, 39, 41–44]). However, this procedure costs the additional computation time. Therefore, it is of particular interest to introduce new approaches, especially for the stations where benchmark methods do not work at all.

To this end, a new machine learning method, the “least absolute shrinkage and selection operator (LASSO)” algorithm, is introduced for downscaling ERA-Interim forecast precipitation data in complex terrain. Compared to standard linear downscaling approaches, LASSO is also well suited for possibly underdetermined linear regression problems, as well as for joint estimation and continuous variable selection. It is tested and validated against three different methods, local intensity scaling (LOCI), quantile-mapping (QM), and stepwise regression (Stepwise), using data from 50 meteorological stations located in the high mountainous region of the central Alps (Figure 1).

This paper is structured as follows. Section 2 describes the ERA-Interim forecast precipitation data and meteorological observations in the study area. Section 3 describes the four downscaling methods as well as the evaluation criteria. The downscaling results and methods comparison are presented in Section 4, while finally a discussion and subsequent conclusions are given in Section 5.

2. Datasets

2.1. ERA-Interim. We make use of the ERA-Interim forecast precipitation data provided by the European Centre for Medium Range Weather Forecast (ECMWF) for the year 1979 onwards and continuing in real time [45, 46]. ERA-Interim shows some improvements when compared with ERA-40 in these aspects: a representation of the hydrological cycle, an improved description of the stratospheric circulation, and an enhanced handling of biases [46–49]. Cycle 31r2 of ECMWF’s Integrated Forecast System (IFS) was used in here. The model in this configuration comprises 60 vertical levels, with the top level at 0.1 hPa; it uses the T255 spectral harmonic representation for the basic dynamical fields and a reduced Gaussian grid (N128) with an approximately uniform spacing of 79 km [46, 50]. ERA-Interim assimilates four analyses per day at 00, 06, 12, and 18 UTC. Furthermore, two 10-day forecasts with a 3-hour resolution are initialized based on the 00:00 UTC and 12:00 UTC analyses.

A variety of data in uniform lat/long grids (0.125°, 0.25°, 0.5°, 0.75°, 1°, 1.125°, 1.5°, 2°, 2.5°, and 3°) are available from ECMWF datasets. Here, we apply 3-hourly forecast precipitation data (03, 06, 09, 12, 15, 18, 21, and 24 UTC) initialized at 00 UTC from 1983–2010 which are projected on the grid of 0.25° × 0.25°. Daily precipitation is calculated from eight times 3-hourly precipitation forecast. In addition, 19 surface and upper-atmosphere variables (Table 1) are applied as predictors according to previous studies (e.g., [38, 40]). These 3-hourly predictors are aggregated into daily averages. The used variables are listed in Table 1. We used 00 UTC initialized forecasts and calculated the elevation in meters a.s.l. by the normalization of the geopotential over the gravity.

2.2. Test Sites. Daily total precipitation of the period 1983–2010 at 50 meteorological stations was made available through the interactive tools of IDAWEb, which is designed by MeteoSwiss (the Swiss Federal Office of Meteorology and Climatology) providing free, available, and extensive archive data of ground level monitoring networks. Table 2 lists the information about stations and Figure 1 shows the locations of test sites. The stations are located within a large range of altitude from 381 m to 3305 m. Among these stations, 16 are located below 500 m and 11 are situated between 500 m and 1000 m, 9 between 1000 m and 1500 m, 8 between 1500 m and 2000 m, and 6 above 2000 m. The observations and ERA-Interim data are processed for the same period. The available data is partitioned into two periods, 1983–1999 for calibration and 1999–2010 for validation. A 1 mm threshold was defined for defining a dry/wet day.

TABLE 1: Predictors from ERA-Interim forecast dataset. All variables are aggregated from 3-hourly to daily averages.

Predictor	Description
P_ERA	ERA-Interim precipitation (mm)
MSLP	Mean sea-level pressure (hPa)
TCW	Total column water (mm)
U_10	10-meter <i>U</i> wind component (m s ⁻¹)
V_10	10-meter <i>V</i> wind component (m s ⁻¹)
WS_10	Wind speed (m s ⁻¹)
FG_10	10-meter wind gust (m s ⁻¹)
LSP	Large-scale precipitation (mm)
H_850	Geopotential height at 850 hPa (m)
H_700	Geopotential height at 700 hPa (m)
H_500	Geopotential height at 500 hPa (m)
T_850	Temperature at 850 hPa (°C)
T_700	Temperature at 700 hPa (°C)
T_500	Temperature at 500 hPa (°C)
RH_850	Relative humidity at 850 hPa (%)
RH_700	Relative humidity at 700 hPa (%)
RH_500	Relative humidity at 500 hPa (%)
SH_850	Specific humidity at 850 hPa (g kg ⁻¹)
SH_700	Specific humidity at 700 hPa (g kg ⁻¹)
SH_500	Specific humidity at 500 hPa (g kg ⁻¹)

It is necessary to note that the data of the stations GUE, PAY, and GVE are used within the ERA-interim data assimilation procedure, given their status as WMO SYNOP stations [46, 51]. According to the information of the ECMWF, it can be assumed that the majority of the stations (47 of 50 sites) are not used by ERA-Interim and therefore represent a fully independent dataset.

3. Methods

3.1. Local Intensity Scaling (LOCI). LOCI is a robust method to directly correct GCM or RCM outputs for local observations. Although GCMs or RCMs are partly unrealistic due to their coarse resolution, they contain valuable information about the actual precipitation [11]. The assumption is realized by a so-called scaling factor, calculated from observation and climate model data of a reference period, which is then expanded to scenarios data. Here, LOCI is applied as the benchmark method for comparison with LASSO. LOCI was developed by Widmann and Bretherton [52]. Widmann et al. [53] used it for scenario precipitation corrections. Not only GCMs, but also RCMs were corrected using the LOCI approach [54–56]. Schmidli et al. [57] further modified LOCI for precipitation occurrence and amount correction, separately. In this study, LOCI is implemented based on a monthly scaling factor which is calculated in three steps as follows:

$$\begin{aligned}
 \text{Fre} \left(P_{\text{ERA}} \geq P_{\text{ERA}}^{\text{thres}} \right) &= \text{Fre} \left(P_{\text{obs}} \geq P_{\text{obs}}^{\text{thres}} \right), \\
 \text{SF} &= \frac{\overline{(P_{\text{obs}} \geq P_{\text{obs}}^{\text{thres}})} - P_{\text{obs}}^{\text{thres}}}{\overline{(P_{\text{ERA}} \geq P_{\text{ERA}}^{\text{thres}})} - P_{\text{ERA}}^{\text{thres}}}, \\
 P_t &= \max \left(P_{\text{obs}}^{\text{thres}} + \text{SF} \left(P_{\text{ERA}}^{\text{Val}} - P_{\text{ERA}}^{\text{thres}} \right), 0 \right),
 \end{aligned} \tag{1}$$

TABLE 2: Test sites information. ERA-Interim grid height is also listed.

Site abbreviation	Site name	Latitude	Longitude	Site elevation (m)	ERA-Interim grid elevation (m)
COV	Piz Corvatsch	46.42	9.82	3305	1618
WFJ	Weissfluhjoch	46.83	9.81	2690	1722
SAE	Säntis	47.25	9.34	2502	1154
GSB	Col du Grand St-Bernard	45.87	7.17	2472	1493
GUE	Gütsch ob Andermatt	46.65	8.62	2287	1432
PIL	Pilatus	46.98	8.25	2106	1154
MLS	Le Moléson	46.55	7.02	1974	1193
SAM	Samedan	46.53	9.88	1709	1661
DOL	La Dôle	46.42	6.10	1670	699
CIM	Cimetta	46.20	8.79	1661	1281
SBE	S. Bernardino	46.46	9.18	1639	1532
ZER	Zermatt	46.03	7.75	1638	1552
CHA	Chasseral	47.13	7.05	1599	669
DAV	Davos	46.81	9.84	1594	1722
MVE	Montana	46.30	7.47	1427	1470
NAP	Napf	47.00	7.93	1403	1115
ULR	Ulrichen	46.50	8.30	1345	1435
SCU	Scuol	46.80	10.28	1303	1818
FRE	Bullet/La Frétaz	46.83	6.58	1205	818
DIS	Disentis/Sedrun	46.70	8.85	1197	1479
ROB	Poschiavo/Robbia	46.35	10.07	1078	1394
ENG	Engelberg	46.82	8.42	1035	1432
CDF	La Chaux-de-Fonds	47.08	6.80	1018	770
PIO	Piotta	46.52	8.68	990	1447
STG	St. Gallen	47.43	9.40	775	1027
VIS	Visp	46.30	7.85	639	1498
RUE	Rünnenberg	47.43	7.88	611	701
FAH	Fahy	47.42	6.93	596	554
INT	Interlaken	46.67	7.87	577	1325
CHU	Chur	46.87	9.53	556	1668
SMA	Zürich/Fluntern	47.38	8.57	555	780
BER	Bern/Zollikofen	46.98	7.47	552	1008
TAE	Aadorf/Tänikon	47.48	8.90	539	897
GLA	Glarus	47.03	9.07	516	1312
PAY	Payerne	46.82	6.95	490	1030
WAE	Wädenswil	47.22	8.68	485	1011
NEU	Neuchâtel	47.00	6.95	485	849
SIO	Sion	46.22	7.33	482	1408
VAD	Vaduz	47.13	9.52	457	1226
PUY	Pully	46.52	6.67	455	1100
CGI	Nyon/Changins	46.40	6.23	455	835
LUZ	Luzern	47.03	8.30	454	1154
REH	Zürich/Affoltern	47.43	8.52	443	780
GUT	Güttingen	47.60	9.28	440	962
SHA	Schaffhausen	47.68	8.62	438	662
ALT	Altdorf	46.88	8.62	438	1193
KLO	Zürich/Kloten	47.48	8.53	426	780
WYN	Wynau	47.25	7.78	422	827
GVE	Genève-Cointrin	46.25	6.13	420	973
AIG	Aigle	46.33	6.92	381	1346

where P_t is the target station precipitation, P_{ERA}^{Val} the un-downscaled ERA-Interim data for validation, P_{ERA}^{thres} ERA-Interim precipitation threshold, and P_{obs}^{thres} the observation threshold and the brackets present the frequency condition judgment function. Here, 1 mm is used to define wet/dry days and SF is the scaling factor. In the first step, an adjusted threshold for ERA-Interim data is found that matches the occurrence of wet/dry days, based on the 1 mm threshold of observation. In a second step, the scaling factor is obtained and then, finally, the target station precipitation is calculated.

3.2. Quantile-Mapping (QM). QM introduced by Panofsky and Brier [58] is a popular statistical transformations approach to correct GCM and RCM outputs straightforwardly [25, 59–65]. The distribution function (e.g., cumulative distribution function, CDF) of model precipitation is first adjusted to match the distribution of observations. Subsequently, this matched distribution is used for unbiased model (or future scenario) data. The mapping is usually implemented based on empirical quantiles or quantiles of gamma distributions [11, 64]. In this study, the corrected ERA-Interim can be obtained via

$$P_t = P_{ERA}^{Val} + \Delta f, \\ \Delta f = f_{obs,cal}^{-1}(f_{ERA,cal}(P_{ERA}^{Val})) - f_{ERA,cal}^{-1}(f_{ERA,cal}(P_{ERA}^{Val})), \quad (2)$$

where $f_{obs,cal}^{-1}$ and $f_{ERA,cal}^{-1}$ is the inverse CDF of observations and ERA-Interim for calibration, respectively, and $f_{ERA,cal}$ is the CDF of P_{ERA}^{Val} .

3.3. Stepwise Regression (Stepwise). Stepwise Regression (Stepwise hereafter) is an automatic procedure where Stepwise combines an ordinary regression (3) with a predictor variable selection procedure. Three main approaches are used in Stepwise according to the relevant selection sequence: forward selection, backward elimination, or bidirectional elimination. The advantage of stepwise regression is easily explained and implemented. Several previous studies have used stepwise regression for different purposes. For example, Harpham and Wilby [39], Hessami et al. [42], and Huth [66] adopted Stepwise for predictor selection; Agnihotri and Mohapatra [30] applied it to occurrence estimation of daily summer monsoon precipitation. In this study, stepwise regression is adopted to test LASSO. Stepwise regression is implemented using backward elimination method with a significance level of 0.05,

$$y_i = x_{ij}\beta_j + \varepsilon, \quad (3)$$

where y is the $i \times 1$ response vector, x is the $i \times j$ variable vector, β is the $j \times 1$ parameter vector, and ε is the random errors. We used the same set of 20 variables (Table 1) for Stepwise and LASSO, for a better method comparison. For precipitation occurrence, y is defined as 1 for wet days (>1 mm) and 0 for dry days (<1 mm). All variables are standardized to make data fall between 0 and 1.

3.4. LASSO Algorithm. Least absolute shrinkage and selection operator (LASSO) is an alternative regularized version of least squares, which is useful for feature selection and to avoid overfitting problems. LASSO shrinks the estimates of the regression coefficients towards zero to prevent overfitting problem and to reduce variables by using a penalty parameter [67, 68]. To simplify understanding, the history of LASSO is introduced briefly. The following equation presents the ordinary least squares regression (OLS) that tries to minimize the error RSS (Root of Sum of Squares),

$$\hat{\beta}_{ols} = \arg \min_{\beta} \left\{ \sum_{i=1}^N \left(y_i - \beta_0 - \sum_{j=1}^v x_{ij}\beta_j \right)^2 \right\}. \quad (4)$$

OLS is not always satisfactory for minimizing the RSS, especially when x contains a large number of variables. A penalty parameter λ was added based on the normal OLS in LASSO (see the following equation):

$$\hat{\beta}_{lasso} = \arg \min_{\beta} \left\{ \sum_{i=1}^N \left(y_i - \beta_0 - \sum_{j=1}^v x_{ij}\beta_j \right)^2 + \lambda \sum_{j=1}^v \beta_j \right\}, \quad (5)$$

where v is the number of variables. LASSO imposes intentionally that some coefficients have to be zero, thus achieving a sparse model. Thus, the penalty parameter (regularizer) λ controls the level of sparsity of the resulting model. In this study, we applied an efficient algorithm for solving LASSO [69]. Also, we defined the value 1 for wet days (>1 mm) and 0 for dry days (<1 mm), and all variables are standardized the same as Stepwise.

3.5. Evaluation Criteria. Precipitation downscaling procedure is implemented in two steps. Firstly, precipitation occurrence is modeled by the four methods, respectively. Secondly, precipitation amount conditional on the occurrence of wet days is modeled subsequently. Please note that the results of precipitation occurrence ranges from 0 to 1. We defined 0.5 as the threshold value to classify dry/wet days. The modeled precipitation amount also could be negative values. Therefore, we set these negative values to zero. The root mean square error (RMSE) and the mean absolute error (MAE) are used for the assessment of precipitation amount, and correspondence ratio (CR) is applied for the evaluation of dry/wet days classification accuracy (6)-(7),

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (P_o - P_m)^2}, \quad (6)$$

$$MAE = \frac{1}{N} \sum_{t=1}^N |P_o - P_m|,$$

where P_o and P_m are observed and modeled precipitation amount on wet days, respectively. N is the number of wet days,

$$CR = \frac{n_{dry} + n_{wet}}{N}, \quad (7)$$

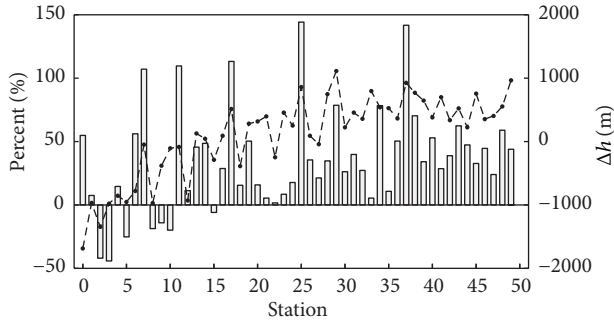


FIGURE 2: Annual precipitation bias (bars) between ERA-Interim and MeteoSwiss stations in percent for the period 1983–2010. Positive values indicate an overestimation of the annual precipitation by ERA-Interim and vice versa. The line with dots indicates altitude differences, Δh , defined as ERA-Interim grid height minus site elevation.

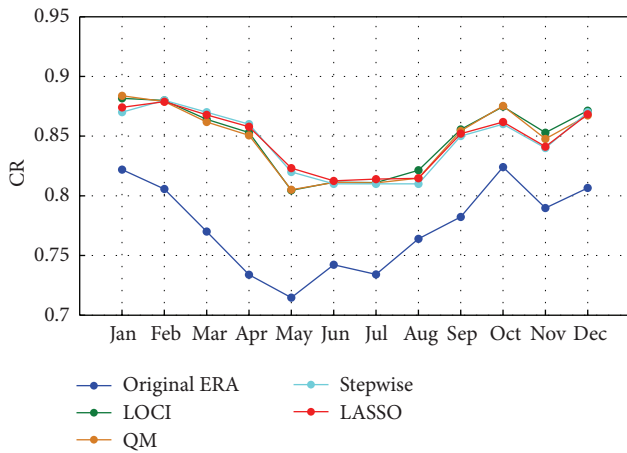


FIGURE 3: Averaged correspondence ratio (CR) for each month in the validation period 1999–2010.

where n_{dry} and n_{wet} are the numbers of dry or wet days that are correctly classified by downscaling methods.

4. Results

4.1. Validation of the Original ERA-Interim Forecast Precipitation Data. Table 3 shows the comparison of ERA-Interim daily precipitation forecasts with observations at 50 meteorological stations from 1983–2010. CR as well as the RMSE and MAE in mm is listed. CR varies from 0.68 to 0.86 and RMSE changes in the range of 4.40–11.13 mm, while MAE ranges from 2.22 to 5.19 mm. The large errors show that there is a great need for the correction and downscaling of ERA-Interim data at local stations. The altitude differences change sharply in a large interval. We use the abbreviations for the characterization of the stations. Station COV is 1687 m higher than ERA-Interim grid height while station SCU is 515 m lower than grid height. ERA-Interim shows good agreement with the occurrence of observations (0.68 to 0.86), but large

deviations with respect to the amount of precipitation on wet days. However, in contrast to temperature estimations, the difference in altitude is unable to explain the observed bias. For example, station DOL has the highest CR with the value of 0.86, but with a large elevation difference of 971 m. Station ZER has a CR of 0.74 but its elevation matches the grid height very well with an altitude gap of only 86 m. This again suggests that the general relationship between precipitation and elevation is not easy to obtain, due to the great variability in the interaction between atmospheric circulation and complex topographical characteristics.

Figure 2 illustrates the bias (in percent) between ERA-Interim data and station precipitation observations for the time period 1983–2010. Positive values indicate that annual precipitation is overestimated by ERA-Interim and vice versa. In general, ERA-Interim data overestimate observations the observed annual precipitation for the majority of test sites; only 7 of 50 sites are underestimated. Among them, station GSB has the largest negative bias with a value of -44.2% . 5 out of 50 sites show an overestimation of more than 100% . Station WFJ and DAV are located in the same ERA-Interim grid. However, ERA-Interim overestimated station WFJ with 7.5% and station DAV with a value of 45.6% . Besides, stations PIL and LUZ are also in the same ERA-Interim grid. ERA-Interim underestimates PIL with 25.2% . In the contrast, ERA-Interim overestimates LUZ with a value of 28.6% . The lower stations that are mainly located at the northern part of the Alps are underestimated by ERA-Interim.

4.2. Evaluation of Downscaling Methods in Precipitation Occurrence. The performance of the four downscaling methods, as well as the ERA-Interim original data in precipitation occurrence in the validation period 1999–2010, is summarized in Table 4. The averaged CR showed that the four downscaling methods were much better than the original ERA-Interim data. QM and LASSO had the same range of CR from 0.82 to 0.88. LOCI ranged from 0.81 to 0.88 and Stepwise varied from 0.82 to 0.87. QM estimated a worse CR than the original ERA data at station GSB. LASSO predicted the worst occurrence at station DOL. The four methods had the same performance at 7 sites. For 11 sites, LASSO and Stepwise were the best methods in precipitation occurrence modeling. Figure 3 illustrates the averaged correspondence ratio for each month in the validation period 1999–2010. It reveals that LOCI, QM, Stepwise, and LASSO captured the general tendency of the monthly precipitation occurrence for the whole year. LASSO and Stepwise identified better dry and wet days than LOCI and QM in May, but they performed worse in October. LASSO classified dry/wet days similarly with Stepwise for all 12 months. The result shows that LASSO is quite suitable for dry/wet days classification.

Figure 4 shows the used variables in Stepwise and LASSO for dry/wet days classification. LASSO selected more variables than Stepwise. For LASSO, P_{ERA} , T_{700} , and TCW were the most frequent (50 sites) variables. RH_{500} was the least frequently (17 sites) applied by LASSO. P_{ERA} and U_{10} were the most frequent (more than 45 sites) variables in Stepwise. T_{500} was the least frequent variable.

TABLE 3: Comparison of ERA-Interim forecast daily precipitation with observations at 50 meteorological stations from 1983 to 2010. CR, RMSE, and MAE in mm are listed. The elevations of sites are different with ERA-Interim grid heights, so the altitude differences (Δh , ERA-Interim grid height minus elevation) are also labeled.

Site	CR	RMSE	MAE	Δh
COV	0.76	4.66	2.37	-1687
WFJ	0.79	5.83	2.93	-968
SAE	0.84	11.13	5.19	-1348
GSB	0.83	10.86	4.59	-979
GUE	0.79	6.96	3.30	-855
PIL	0.82	9.18	3.97	-952
MLS	0.79	5.40	2.73	-781
SAM	0.68	4.84	2.62	-48
DOL	0.86	6.04	2.76	-971
CIM	0.76	8.97	3.45	-380
SBE	0.74	9.16	3.91	-107
ZER	0.74	5.33	2.64	-86
CHA	0.80	5.23	2.56	-930
DAV	0.76	5.10	2.69	128
MVE	0.75	5.52	2.75	43
NAP	0.82	5.70	2.79	-288
ULR	0.75	6.33	3.14	90
SCU	0.69	4.98	2.74	515
FRE	0.82	5.08	2.46	-387
DIS	0.75	6.12	3.05	282
ROB	0.77	5.46	2.50	316
ENG	0.81	5.91	3.01	397
CDF	0.84	4.87	2.33	-248
PIO	0.72	8.17	3.58	457
STG	0.81	5.48	2.74	252
VIS	0.68	5.66	2.98	859
RUE	0.79	4.62	2.38	90
FAH	0.81	4.40	2.22	-42
INT	0.79	5.24	2.70	748
CHU	0.72	5.19	2.81	1112
SMA	0.80	4.85	2.44	225
BER	0.78	5.07	2.59	456
TAE	0.80	4.91	2.55	358
GLA	0.80	5.55	2.87	796
PAY	0.77	5.45	2.84	540
WAE	0.81	5.55	2.71	526
NEU	0.77	5.08	2.59	364
SIO	0.69	5.18	2.80	926
VAD	0.76	5.11	2.69	769
PUY	0.80	5.76	2.79	645
CGI	0.79	4.99	2.56	380
LUZ	0.78	5.46	2.79	700
REH	0.79	4.73	2.42	337
GUT	0.77	5.01	2.66	522
SHA	0.78	4.47	2.33	224
ALT	0.77	5.32	2.74	755
KLO	0.78	4.83	2.49	354
WYN	0.79	4.71	2.37	405
GVE	0.78	5.13	2.61	553
AIG	0.79	4.82	2.44	965
Average	0.78	5.79	2.84	100

TABLE 4: Comparison of downscaling methods, as well as the original ERA-Interim forecast data in precipitation occurrence in the validation period 1999–2010.

Site	ERA	LOCI	CR		LASSO
			QM	Stepwise	
COV	0.76	0.85	0.84	0.82	0.82
WFJ	0.78	0.82	0.82	0.85	0.85
SAE	0.83	0.84	0.84	0.84	0.84
GSB	0.83	0.83	0.82	0.85	0.85
GUE	0.78	0.81	0.82	0.83	0.83
PIL	0.81	0.84	0.84	0.83	0.83
MLS	0.76	0.83	0.83	0.84	0.84
SAM	0.68	0.86	0.85	0.84	0.84
DOL	0.86	0.88	0.88	0.86	0.85
CIM	0.77	0.85	0.84	0.87	0.87
SBE	0.74	0.83	0.82	0.86	0.86
ZER	0.74	0.84	0.84	0.84	0.84
CHA	0.80	0.85	0.85	0.83	0.83
DAV	0.75	0.84	0.84	0.85	0.84
MVE	0.74	0.83	0.83	0.85	0.85
NAP	0.81	0.85	0.85	0.84	0.84
ULR	0.75	0.84	0.83	0.84	0.84
SCU	0.69	0.86	0.86	0.85	0.85
FRE	0.82	0.87	0.87	0.85	0.85
DIS	0.75	0.84	0.84	0.83	0.83
ROB	0.76	0.84	0.83	0.86	0.86
ENG	0.80	0.82	0.82	0.84	0.84
CDF	0.84	0.86	0.87	0.85	0.85
PIO	0.71	0.84	0.83	0.87	0.87
STG	0.81	0.85	0.84	0.84	0.84
VIS	0.68	0.85	0.85	0.85	0.85
RUE	0.78	0.85	0.84	0.83	0.83
FAH	0.81	0.86	0.86	0.85	0.85
INT	0.79	0.85	0.85	0.85	0.85
CHU	0.71	0.84	0.83	0.84	0.84
SMA	0.80	0.86	0.86	0.83	0.83
BER	0.77	0.86	0.86	0.85	0.85
TAE	0.80	0.86	0.86	0.85	0.85
GLA	0.79	0.84	0.84	0.86	0.86
PAY	0.76	0.86	0.85	0.85	0.85
WAE	0.81	0.85	0.85	0.84	0.85
NEU	0.77	0.87	0.86	0.85	0.85
SIO	0.69	0.87	0.87	0.87	0.88
VAD	0.76	0.86	0.85	0.85	0.85
PUY	0.79	0.87	0.86	0.86	0.86
CGI	0.78	0.88	0.88	0.87	0.87
LUZ	0.77	0.84	0.84	0.84	0.84
REH	0.79	0.85	0.85	0.83	0.84
GUT	0.78	0.84	0.84	0.84	0.84
SHA	0.78	0.85	0.85	0.83	0.83
ALT	0.77	0.84	0.84	0.84	0.84
KLO	0.79	0.85	0.85	0.84	0.84
WYN	0.79	0.86	0.86	0.85	0.85
GVE	0.78	0.88	0.87	0.86	0.86
AIG	0.78	0.86	0.86	0.87	0.86
Average	0.77	0.85	0.85	0.85	0.85

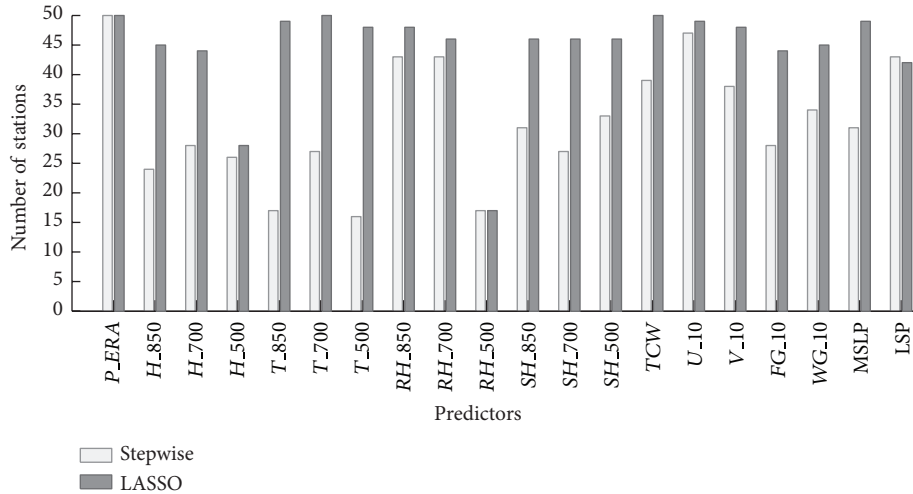


FIGURE 4: Selected variables in Stepwise and LASSO for dry/wet days classification.

For the individual sites, station INT selected least variables (12 variables) while station COV, CHA, and CGI chose the most variables (20 variables) in LASSO. Stepwise selected 16 variables at stations GSB, CHA, FAH, and SMA, while it only applied 9 variables at stations SAM and PIO. LASSO applied more variables compared with Stepwise for the majority of stations (49 stations). Although, different variables are applied by the two methods, the performances of dry/wet days classification were similar.

4.3. Evaluation of Downscaling Methods in Precipitation Amount. The overall performance of four downscaling methods, as well as the ERA-Interim original data in precipitation amount conditional on the occurrence of a wet day for the validation period 1999–2010, is summarized in Table 5. On the average, Stepwise and LASSO performed similarly in RMSE and MAE, which were much smaller than LOCI and QM. The reduction of error of Stepwise was 12.2% of RMSE and 15.7% of MAE, respectively. LASSO was slightly better than Stepwise in MAE (16.3%) and it had the same error reduction with Stepwise in RMSE (12.2%), whereas LOCI and QM were the worst methods. The reduction of RMSE was only 3.5% and 0.9% for LOCI and QM, respectively. In total, LOCI and QM did not reduce RMSE at 13 stations such as station WFJ. QM was the worst method at 24 stations while LOCI performed the worst in RMSE at stations ENG and CDF. LASSO outperformed the other methods at 9 stations. Stepwise reduced the most RMSE at 40 stations compared with other three methods. However, the four downscaling methods failed to reduce the errors at station CHA. In terms of MAE, LOCI outperformed the other methods at 26 stations. LASSO reduced the most MAE at 22 stations. QM method was not able to reduce MAE at 5 stations. It could be concluded that LOCI and QM method are not always suitable for local stations. It has to be noted that the four downscaling methods did not work for MAE reduction at station DOL,

which has a large elevation difference of 971 m against ERA-Interim grid height.

Taking station WFJ as an example, Figures 5 and 6 illustrate the downscaled daily precipitation for the four downscaling methods, as well as the original ERA-Interim data in January and July, respectively, for the validation period 1999–2010. In general, four downscaling methods capture the tendency of the daily precipitation in January and July. However, LOCI and QM are generally inclined to misidentify the heavy precipitation events, in particular in January 2008 and July 2003. Stepwise and LASSO underestimated heavy precipitation events compared with LOCI and QM.

Figure 7 illustrates the comparison of observations with downscaled annual precipitation for the four methods, as well as the original ERA-Interim data in the validation period 1999–2010. LASSO and Stepwise predicted the annual precipitation best. In general, LOCI and QM modeled drier results, with mean value of -7.7% and -5.0% , respectively. 31 out of 50 sites were underestimated by LASSO with the range from -19% (station KLO) to -0.1% (station RUE) while 19 sites were overestimated by 0.2% (station SCU) to 31.6% (station MLS). The majority of sites were underestimated by LOCI (48 out of 50 sites) and QM (40 out of 50 sites) while 30 of 50 sites were underestimated by Stepwise. Figure 8 compares the averaged sum of daily precipitation for each month between observations and the four downscaling methods, as well as the original ERA-Interim data in the validation period 1999–2010. It is easy to find that LASSO and Stepwise predicted more precipitation in the whole year than LOCI and QM with exception of August and November. LASSO reproduced best monthly precipitation for February, March, April, and November. Stepwise was the best method for May, June, July, and October. QM performed best in January, August, and September. LOCI outperformed the other methods in December.

TABLE 5: Comparison of observations with downscaled daily precipitation for the four downscaling methods, as well as the original ERA-Interim data in precipitation amount (mm) on wet days for the validation period 1999–2010.

Site	RMSE					MAE				
	ERA	LOCI	QM	Stepwise	LASSO	ERA	LOCI	QM	Stepwise	LASSO
COV	4.53	4.14	4.31	3.87	3.87	2.33	1.76	1.83	1.80	1.78
WFJ	6.12	6.29	6.45	5.47	5.46	3.03	2.89	2.95	2.71	2.68
SAE	11.89	11.09	11.26	9.74	9.75	5.56	5.49	5.58	5.01	5.01
GSB	11.20	10.59	10.65	9.18	9.18	4.61	4.72	4.78	4.49	4.48
GUE	6.50	6.38	6.79	5.85	5.85	3.16	2.84	2.91	2.95	2.94
PIL	10.42	10.04	9.90	9.57	9.56	4.30	4.21	4.18	4.23	4.20
MLS	5.85	5.28	5.65	4.67	4.66	3.02	2.41	2.50	2.37	2.34
SAM	4.73	3.90	4.05	3.60	3.60	2.58	1.46	1.54	1.55	1.53
DOL	5.71	5.92	6.02	5.40	5.39	2.63	2.64	2.70	2.72	2.70
CIM	8.79	9.46	9.56	8.09	8.09	3.48	3.46	3.52	3.52	3.52
SBE	9.27	9.08	9.91	7.51	7.50	3.98	3.76	3.98	3.52	3.51
ZER	5.10	4.16	4.50	3.65	3.65	2.55	1.61	1.65	1.64	1.62
CHA	5.50	5.61	5.77	5.52	5.52	2.62	2.47	2.52	2.60	2.58
DAV	5.18	4.82	4.89	4.44	4.45	2.72	2.11	2.15	2.09	2.07
MVE	5.32	4.86	4.97	4.21	4.21	2.72	2.08	2.11	2.14	2.13
NAP	6.15	6.47	6.50	6.10	6.10	2.97	2.98	3.00	2.95	2.94
ULR	6.23	5.92	6.21	5.68	5.71	3.11	2.52	2.60	2.70	2.69
SCU	4.93	3.72	3.98	3.55	3.55	2.72	1.46	1.54	1.57	1.54
FRE	5.15	5.43	5.56	4.91	4.91	2.46	2.30	2.34	2.26	2.25
DIS	5.97	5.33	5.49	5.29	5.29	3.00	2.18	2.20	2.37	2.36
ROB	5.72	6.14	6.24	5.09	5.08	2.66	2.40	2.47	2.37	2.36
ENG	6.16	6.64	6.41	5.79	5.79	3.13	3.17	3.10	2.93	2.90
CDF	5.01	5.44	5.44	4.89	4.89	2.34	2.37	2.40	2.36	2.33
PIO	7.88	7.72	7.98	6.76	6.77	3.54	2.99	3.09	3.16	3.16
STG	5.67	5.81	5.92	5.55	5.53	2.76	2.54	2.58	2.60	2.57
VIS	5.49	4.12	4.53	3.75	3.75	2.92	1.52	1.60	1.62	1.61
RUE	4.62	4.57	4.68	4.15	4.15	2.40	1.98	2.03	2.00	1.98
FAH	4.67	4.69	4.78	4.44	4.45	2.28	1.99	2.04	2.05	2.02
INT	5.21	5.30	5.34	4.82	4.81	2.71	2.33	2.35	2.33	2.32
CHU	5.19	4.48	4.82	4.20	4.21	2.80	1.88	1.96	1.91	1.90
SMA	4.86	4.95	5.14	4.71	4.72	2.44	2.17	2.23	2.18	2.16
BER	5.14	5.16	5.23	4.68	4.68	2.64	2.15	2.20	2.18	2.16
TAE	4.74	4.80	4.97	4.51	4.51	2.54	2.24	2.30	2.28	2.26
GLA	5.60	6.02	6.05	5.29	5.29	2.87	2.80	2.82	2.65	2.64
PAY	5.30	4.51	4.60	3.96	3.97	2.83	1.86	1.92	1.90	1.88
WAE	5.50	5.96	6.05	5.40	5.40	2.74	2.67	2.71	2.67	2.64
NEU	5.00	4.85	4.90	4.34	4.34	2.60	1.95	2.00	1.97	1.94
SIO	4.98	3.52	3.65	3.09	3.08	2.74	1.37	1.43	1.48	1.46
VAD	5.26	4.61	4.83	4.38	4.38	2.72	1.87	1.94	1.91	1.90
PUY	5.58	5.57	5.76	5.06	5.06	2.76	2.28	2.36	2.44	2.43
CGI	4.95	4.50	4.69	4.14	4.15	2.55	1.83	1.90	1.86	1.85
LUZ	5.40	5.48	5.52	5.21	5.21	2.81	2.45	2.48	2.55	2.52
REH	4.91	4.82	4.88	4.54	4.55	2.46	2.04	2.07	2.04	2.03
GUT	4.84	4.34	4.56	3.96	3.97	2.60	1.91	1.98	1.91	1.89
SHA	4.36	4.25	4.34	3.90	3.90	2.31	1.85	1.88	1.83	1.80
ALT	5.47	5.47	5.57	5.23	5.23	2.80	2.39	2.44	2.40	2.37
KLO	4.90	4.83	4.97	4.62	4.63	2.51	2.07	2.11	2.07	2.05
WYN	4.87	5.23	5.31	4.60	4.60	2.46	2.21	2.24	2.19	2.17
GVE	5.13	4.72	4.90	4.53	4.53	2.61	1.87	1.96	1.93	1.93
AIG	4.69	4.55	4.65	4.02	4.02	2.43	1.91	1.97	1.99	1.97
Average	5.83	5.63	5.78	5.12	5.12	2.87	2.41	2.46	2.42	2.40

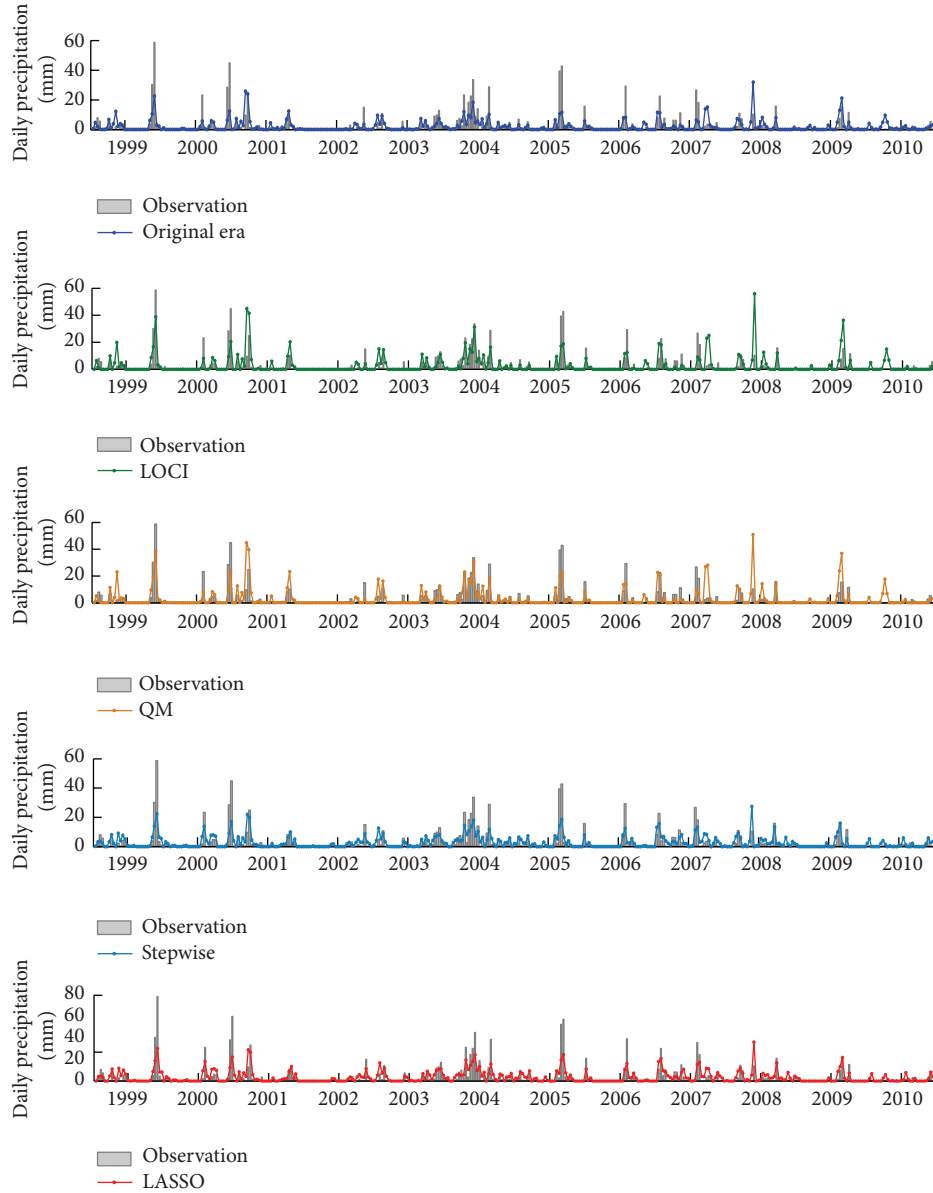


FIGURE 5: Comparison of observations with downscaled daily precipitation for the four downsampling methods, as well as the original ERA-Interim data in January at station WFJ during the validation period 1999–2010

Figure 9 shows the used variables in Stepwise and LASSO for precipitation amount prediction on wet days. LASSO again selected more variables than Stepwise. For LASSO, P_ERA , H_850 , H_700 , H_500 , T_850 , and $MSLP$ were applied at all 50 sites. SH_500 was applied at 25 sites. P_ERA was the sole variable used by all test sites by Stepwise. SH_850 and U_10 were the second most frequent (34 sites) variables. RH_700 and RH_500 were the least frequent variables (10 sites) by Stepwise. For the individual sites, station VAD used the least variables (11 variables) while 19 stations selected all 20 variables in LASSO. Station SBE applied the most variables (16 out of 20 variables) while station WYN only applied 4 variables in Stepwise.

5. Discussion and Conclusion

The comparison between ERA-Interim and observations showed that ERA-Interim has a large error (5.79 mm of RMSE and 2.84 mm of MAE) in the central Alps (Table 3). Thus, there is a great need for the correction and downscaling of ERA-Interim data. This study compared four downsampling methods, LOCI, QM, Stepwise, and LASSO, for downsampling of ERA-Interim daily precipitation data in the central Alps.

As a frequent input variable for hydrological models, daily precipitation is always corrected or downscaled due to the limits of rain gauges. In the previous studies (e.g., [57, 63]), LOCI and QM methods have been widely used for

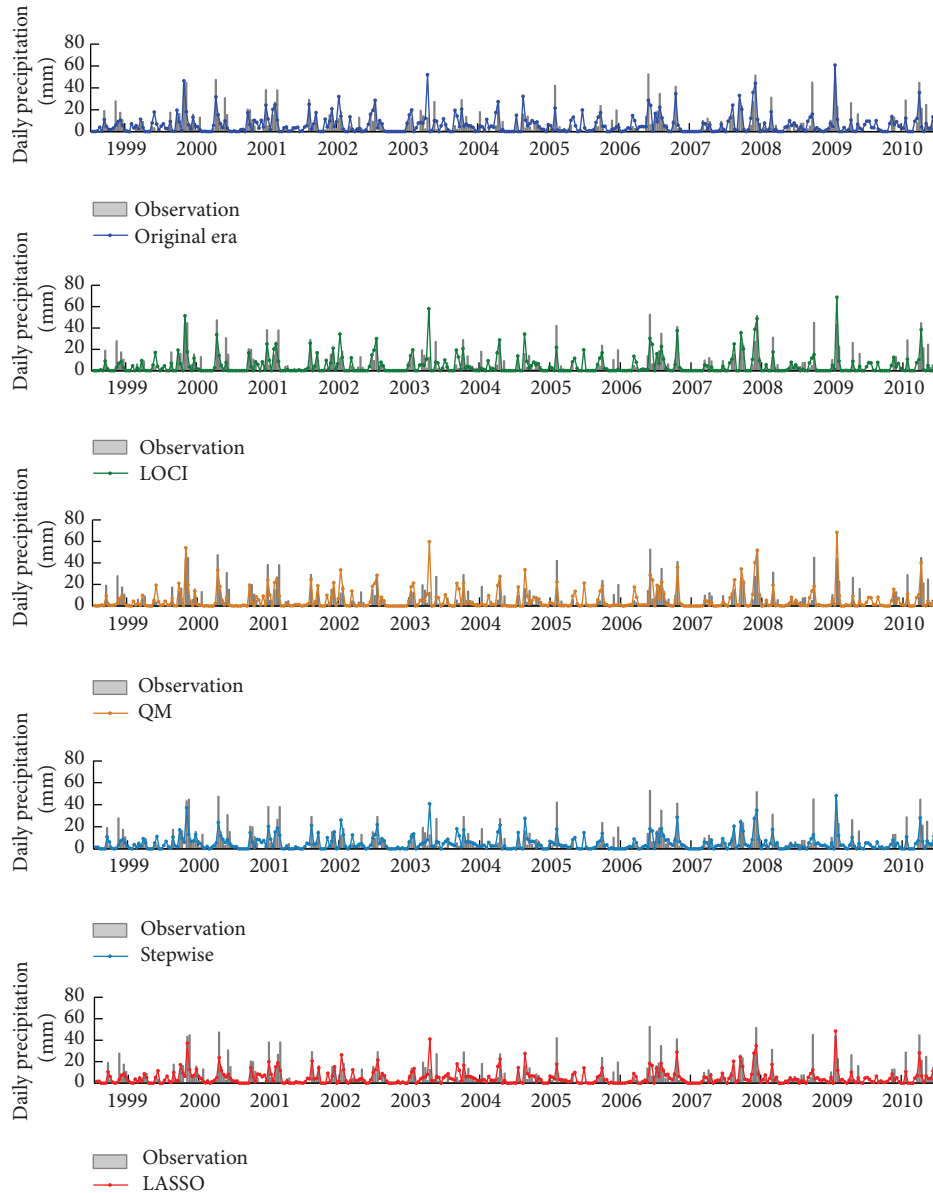


FIGURE 6: Comparison of observations with downscaled daily precipitation for the four downsampling methods, as well as the original ERA-Interim data in July at station WFJ during the validation period 1999–2010.

the bias correction with the advantages of maintaining the variation and distribution of historical data. Although LOCI and QM captured the best estimation of daily precipitation occurrence, the reduction of error is not significant by these two methods and even worse than the original ERA-Interim data for 13 out of 50 sites. It demonstrates that straightforward methods are not always suitable for downscaling local observations in complex terrain.

LASSO algorithm simulated the occurrence of daily precipitation as well as LOCI, QM, and Stepwise; it captured the occurrence for all test sites generally. Compared to the other three downsampling methods, it reduced the most amount error to 16.3% of MAE and 12.2% RMSE. LASSO

also predicted the best annual and monthly precipitation. Although the LASSO algorithm has been developed for more than 15 years by statistician, the application in geosciences is still at the early-stage (e.g., [70]). A main practical challenge in applying LASSO for precipitation downscaling is that precipitation in heterogeneous terrain is such a complex process which tends to use more variables and to overestimate observations. It has shown that LASSO tends to use more variables than Stepwise. To avoid the overfitting problem, the penalty (or regularization) parameter λ plays a key role. The cross-validation and hundreds of runs are necessary to find an appropriate penalty parameter. Therefore, LASSO is a little bit time-consuming than the other three methods.

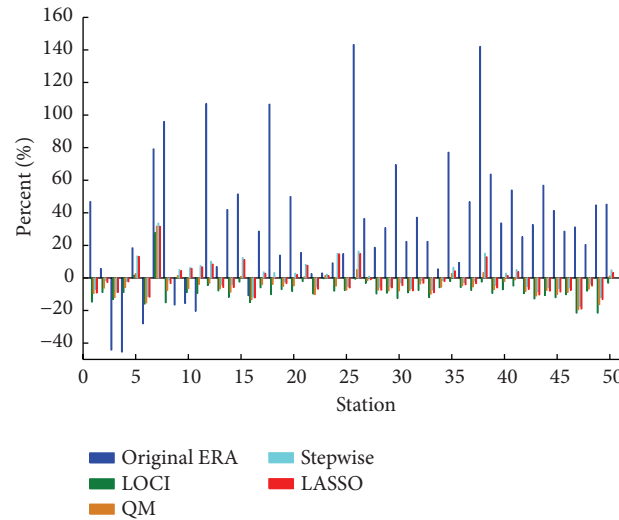


FIGURE 7: Percentage of annual precipitation bias between observations and the four downscaled results, as well as the original ERA-Interim data in the validation period 1999–2010.

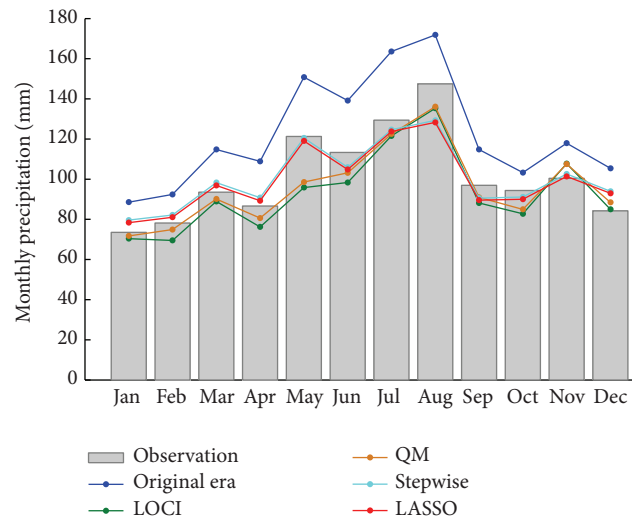


FIGURE 8: Comparison of averaged sum of daily precipitation for each month between observations and the four downscaled results, as well as the original ERA-Interim data in the validation period 1999–2010.

So far, the work presented herein has been limited to the central Alps with 50 meteorological stations providing calibration/validation data sets for testing LASSO algorithm. It would be necessary to extend LASSO method to other high mountainous areas around the world. Besides, this study focuses on the local stations at present; it is of special interest to extend LASSO method for nonstation areas. A limited number of variables (20 variables in this study) derived from ERA-Interim data were applied for LASSO; however, it should also be investigated whether other potential variables such as vapor pressure can be used in the presented approach.

This study focused on the daily total precipitation. Certainly, higher temporal resolution data, such as 3-hourly, would be of great interest in further investigations. Furthermore, applying other reanalysis data sets based on different land surface representations could also be valuable for validating LASSO algorithm.

Conflict of Interests

The authors declare that there is no conflict of interests in this paper.

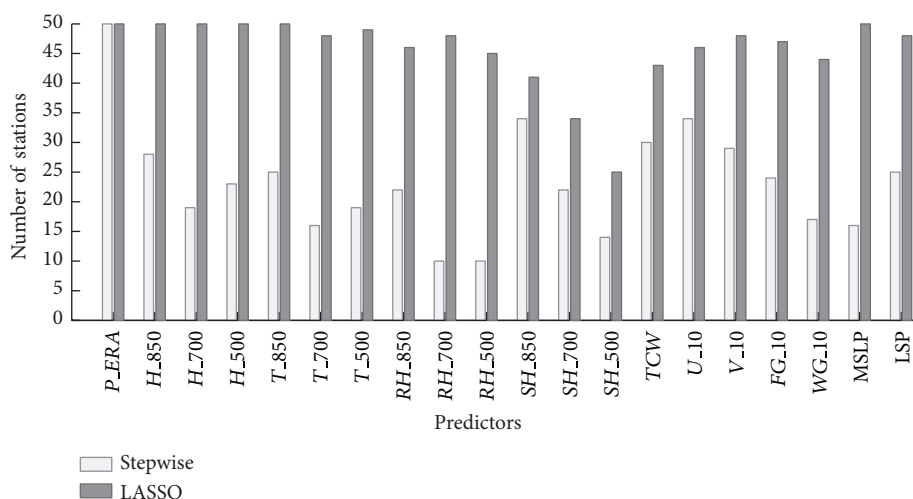


FIGURE 9: Selected variables in Stepwise and LASSO for precipitation amount prediction on wet days.

Acknowledgments

This research has been funded by China Scholarship Council (CSC). ERA-Interim data was supported by the ECMWF. The authors would like to thank MeteoSwiss (the Swiss Federal Office of Meteorology and Climatology) for providing the needed meteorological data. The authors appreciate OURANOS and Josef Schmid (Ludwig-Maximilians-Universität München) for providing a tool for the data processing.

References

- [1] R. E. Benestad, I. Hanssen-Bauer, and E. J. Førland, "An evaluation of statistical models for downscaling precipitation and their ability to capture long-term trends," *International Journal of Climatology*, vol. 27, no. 5, pp. 649–665, 2007.
- [2] X. Beuchat, B. Schaefli, M. Soutter, and A. Mermoud, "Toward a robust method for subdaily rainfall downscaling from daily data," *Water Resources Research*, vol. 47, no. 9, 2011.
- [3] H. Guan, J. L. Wilson, and H. Xie, "A cluster-optimizing regression-based approach for precipitation spatial downscaling in mountainous terrain," *Journal of Hydrology*, vol. 375, no. 3–4, pp. 578–588, 2009.
- [4] M. Bray, D. Han, Y. Xuan, P. Bates, and M. Williams, "Rainfall uncertainty for extreme events in NWP downscaling model," *Hydrological Processes*, vol. 25, no. 9, pp. 1397–1406, 2011.
- [5] R. E. Chandler and H. S. Wheeler, "Analysis of rainfall variability using generalized linear models: a case study from the west of Ireland," *Water Resources Research*, vol. 38, no. 10, pp. 10–1–10–11, 2002.
- [6] B. L. Lamprey, "Comparison of gridded multisatellite rainfall estimates with gridded gauge rainfall over West Africa," *Journal of Applied Meteorology and Climatology*, vol. 47, no. 1, pp. 185–205, 2008.
- [7] C. Daly, M. Halbleib, J. I. Smith et al., "Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States," *International Journal of Climatology*, vol. 28, no. 15, pp. 2031–2064, 2008.
- [8] R. Fealy and J. Sweeney, "Statistical downscaling of precipitation for a selection of sites in Ireland employing a generalised linear modelling approach," *International Journal of Climatology*, vol. 27, no. 15, pp. 2083–2094, 2007.
- [9] L. Ferraris, S. Gabellani, N. Rebora, and A. Provenzale, "A comparison of stochastic models for spatial rainfall downscaling," *Water Resources Research*, vol. 39, no. 12, 2003.
- [10] A. M. Anders, G. H. Roe, D. R. Durran, and J. R. Minder, "Small-scale spatial gradients in climatological precipitation on the olympic Peninsula," *Journal of Hydrometeorology*, vol. 8, no. 5, pp. 1068–1081, 2007.
- [11] D. Maraun, F. Wetterhall, A. M. Ireson et al., "Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user," *Reviews of Geophysics*, vol. 48, no. 3, pp. 1–34, 2010.
- [12] J. M. Eden, M. Widmann, D. Grawe, and S. Rast, "Skill, correction, and downscaling of GCM-simulated precipitation," *Journal of Climate*, vol. 25, no. 11, pp. 3970–3984, 2012.
- [13] M. G. Grillakis, A. G. Koutroulis, and I. K. Tsanis, "Multi-segment statistical bias correction of daily GCM precipitation output," *Journal of Geophysical Research D: Atmospheres*, vol. 118, no. 8, pp. 3150–3162, 2013.
- [14] T. Lafon, S. Dadson, G. Buys, and C. Prudhomme, "Bias correction of daily precipitation simulated by a regional climate model: a comparison of methods," *International Journal of Climatology*, vol. 33, no. 6, pp. 1367–1381, 2013.
- [15] E. P. Maurer, A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen, "A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States," *Journal of Climate*, vol. 15, no. 22, pp. 3237–3251, 2002.
- [16] B. Ahrens, "Rainfall downscaling in an alpine watershed applying a multiresolution approach," *Journal of Geophysical Research: Atmospheres*, vol. 108, no. D8, 2003.
- [17] S. P. Charles, B. C. Bates, I. N. Smith, and J. P. Hughes, "Statistical downscaling of daily precipitation from observed and modelled atmospheric fields," *Hydrological Processes*, vol. 18, no. 8, pp. 1373–1394, 2004.
- [18] F. Giorgi and L. O. Mearns, "Approaches to the simulation of regional climate change: a review," *Reviews of Geophysics*, vol. 29, no. 2, pp. 191–216, 1991.
- [19] M. J. Muerth, B. Gauvin St-Denis, S. Ricard et al., "On the need for bias correction in regional climate scenarios to assess climate

- change impacts on river runoff," *Hydrology and Earth System Sciences*, vol. 17, no. 3, pp. 1189–1204, 2013.
- [20] C. Xu, "From GCMs to river flow: a review of downscaling methods and hydrologic modelling approaches," *Progress in Physical Geography*, vol. 23, no. 2, pp. 229–249, 1999.
- [21] M. Bernhardt, K. Schulz, G. E. Liston, and G. Zängl, "The influence of lateral snow redistribution processes on snow melt and sublimation in alpine regions," *Journal of Hydrology*, vol. 424–425, pp. 196–206, 2012.
- [22] K. F. Ahmed, G. Wang, J. Silander et al., "Statistical downscaling and bias correction of climate model outputs for climate change impact assessment in the U.S. northeast," *Global and Planetary Change*, vol. 100, pp. 320–332, 2013.
- [23] T. Crawford, N. L. Betts, and D. Favis-Mortlock, "GCM grid-box choice and predictor selection associated with statistical downscaling of daily precipitation over Northern Ireland," *Climate Research*, vol. 34, no. 2, pp. 145–160, 2007.
- [24] L. Gao, M. Bernhardt, and K. Schulz, "Elevation correction of ERA-Interim temperature data in complex terrain," *Hydrology and Earth System Sciences*, vol. 16, no. 12, pp. 4661–4673, 2012.
- [25] L. Gudmundsson, J. B. Bremnes, J. E. Haugen, and T. Engen-Skaugen, "Technical note: downscaling RCM precipitation to the station scale using statistical transformations—a comparison of methods," *Hydrology and Earth System Sciences*, vol. 16, no. 9, pp. 3383–3390, 2012.
- [26] E. Hertig and J. Jacobeit, "Assessments of Mediterranean precipitation changes for the 21st century using statistical downscaling techniques," *International Journal of Climatology*, vol. 28, no. 8, pp. 1025–1045, 2008.
- [27] R. L. Wilby and T. M. L. Wigley, "Downscaling general circulation model output: a review of methods and limitations," *Progress in Physical Geography*, vol. 21, no. 4, pp. 530–548, 1997.
- [28] E. Zorita and H. Von Storch, "The analog method as a simple statistical downscaling technique: comparison with more complicated methods," *Journal of Climate*, vol. 12, no. 8, pp. 2474–2489, 1999.
- [29] R. L. Wilby, C. W. Dawson, and E. M. Barrow, "SDSM—a decision support tool for the assessment of regional climate change impacts," *Environmental Modelling and Software*, vol. 17, no. 2, pp. 147–159, 2002.
- [30] G. Agnihotri and M. Mohapatra, "Prediction of occurrence of daily summer monsoon precipitation over Karnataka," *Meteorological Applications*, vol. 19, no. 2, pp. 130–139, 2012.
- [31] R. E. Benestad, "A comparison between two empirical downscaling strategies," *International Journal of Climatology*, vol. 21, no. 13, pp. 1645–1668, 2001.
- [32] G. Bürger and Y. Chen, "Regression-based downscaling of spatial variability for hydrologic applications," *Journal of Hydrology*, vol. 311, no. 1–4, pp. 299–317, 2005.
- [33] E. Hertig and J. Jacobeit, "A novel approach to statistical downscaling considering nonstationarities: application to daily precipitation in the Mediterranean area," *Journal of Geophysical Research-Atmospheres*, vol. 118, 2013.
- [34] J. T. Schoof and S. C. Pryor, "Downscaling temperature and precipitation: a comparison of regression-based methods and artificial neural networks," *International Journal of Climatology*, vol. 21, no. 7, pp. 773–790, 2001.
- [35] H. Tatli, H. N. Dalfes, and Ş. S. Menteş, "A statistical downscaling method for monthly total precipitation over Turkey," *International Journal of Climatology*, vol. 24, no. 2, pp. 161–180, 2004.
- [36] A. Bardossy, I. Bogardi, and I. Matyasovszky, "Fuzzy rule-based downscaling of precipitation," *Theoretical and Applied Climatology*, vol. 82, no. 1–2, pp. 119–129, 2005.
- [37] A. J. Cannon, "Nonlinear analog predictor analysis: A coupled neural network/analog model for climate downscaling," *Neural Networks*, vol. 20, no. 4, pp. 444–453, 2007.
- [38] S. T. Chen, P. S. Yu, and Y. H. Tang, "Statistical downscaling of daily precipitation using support vector machines and multivariate analysis," *Journal of Hydrology*, vol. 385, no. 1–4, pp. 13–22, 2010.
- [39] C. Harpham and R. L. Wilby, "Multi-site downscaling of heavy daily precipitation occurrence and amounts," *Journal of Hydrology*, vol. 312, no. 1–4, pp. 235–255, 2005.
- [40] D. N. Kumar, A. Anandhi, V. V. Srinivas, and R. S. Nanjundiah, "Downscaling precipitation to river basin in India for IPCC SRES scenarios using support vector machine," *International Journal of Climatology*, vol. 28, no. 3, pp. 401–420, 2008.
- [41] S. Ghosh, "SVM-PGSL coupled approach for statistical downscaling to predict rainfall from GCM output," *Journal of Geophysical Research D: Atmospheres*, vol. 115, no. 22, Article ID D22102, 2010.
- [42] M. Hessami, P. Gachon, T. B. M. J. Ouarda, and A. St-Hilaire, "Automated regression-based statistical downscaling tool," *Environmental Modelling and Software*, vol. 23, no. 6, pp. 813–834, 2008.
- [43] J. Schmidli, C. Frei, and C. Schär, "Reconstruction of mesoscale precipitation fields from sparse observations in complex terrain," *Journal of Climate*, vol. 14, no. 15, pp. 3289–3306, 2001.
- [44] R. L. Wilby and T. M. L. Wigley, "Precipitation predictors for downscaling: observed and general circulation model relationships," *International Journal of Climatology*, vol. 20, no. 6, pp. 641–661, 2000.
- [45] P. Berrisford, D. Dee, K. Fielding et al., *The ERA-Interim Archive (Version 1.0)*, ERA Report Series, European Centre for Medium Range Weather Forecasts, Berkshire, UK, 2009.
- [46] D. P. Dee, S. M. Uppala, A. J. Simmons et al., "The ERA-Interim reanalysis: Configuration and performance of the data assimilation system," *Quarterly Journal of the Royal Meteorological Society*, vol. 137, no. 656, pp. 553–597, 2011.
- [47] D. P. Dee and S. Uppala, "Variational bias correction of satellite radiance data in the ERA-Interim reanalysis," *Quarterly Journal of the Royal Meteorological Society*, vol. 135, no. 644, pp. 1830–1841, 2009.
- [48] A. Simmons, S. Uppala, D. Dee et al., "ERA-Interim: new ECMWF reanalysis products from 1989 onwards," *ECMWF Newsletter*, vol. 110, pp. 25–35, 2006.
- [49] D. Dee and S. Uppala, "Variational bias correction in ERA-Interim," ECMWF Technical Memorandum 575, 2008.
- [50] S. Uppala, D. Dee, and S. Kobayashi, "Towards a climate data assimilation system: status update of ERA-Interim," *ECMWF Newsletter*, vol. 115, pp. 12–18, 2008.
- [51] A. J. Simmons, K. M. Willett, P. D. Jones, P. W. Thorne, and D. P. Dee, "Low-frequency variations in surface atmospheric humidity, temperature, and precipitation: inferences from reanalyses and monthly gridded observational data sets," *Journal of Geophysical Research D: Atmospheres*, vol. 115, no. 1, Article ID D01110, 2010.
- [52] M. Widmann and C. S. Bretherton, "Validation of mesoscale precipitation in the NCEP reanalysis using a new gridcell dataset for the Northwestern United States," *Journal of Climate*, vol. 13, no. 11, pp. 1936–1950, 2000.

- [53] M. Widmann, C. S. Bretherton, and E. P. Salathé Jr., "Statistical precipitation downscaling over the northwestern united states using numerically simulated precipitation as a predictor," *Journal of Climate*, vol. 16, no. 5, pp. 799–816, 2003.
- [54] T. Engen-Skaugen, "Refinement of dynamically downscaled precipitation and temperature scenarios," *Climatic Change*, vol. 84, no. 3-4, pp. 365–382, 2007.
- [55] L. P. Graham, J. Andréasson, and B. Carlsson, "Assessing climate change impacts on hydrology from an ensemble of regional climate models, model scales and linking methods: a case study on the Lule River basin," *Climatic Change*, vol. 81, no. 1, pp. 293–307, 2007.
- [56] R. Leander, T. A. Buishand, B. J. J. M. van den Hurk, and M. J. M. de Wit, "Estimated changes in flood quantiles of the river Meuse from resampling of regional climate model output," *Journal of Hydrology*, vol. 351, no. 3-4, pp. 331–343, 2008.
- [57] J. Schmidli, C. Frei, and P. L. Vidale, "Downscaling from GCM precipitation: a benchmark for dynamical and statistical downscaling methods," *International Journal of Climatology*, vol. 26, no. 5, pp. 679–689, 2006.
- [58] H. W. Panofsky and G. W. Brier, *Some Applications of Statistics to Meteorology*, The Pennsylvania State University Press, University Park, Pa, USA, 1968.
- [59] J. Boé, L. Terray, F. Habets, and E. Martin, "Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies," *International Journal of Climatology*, vol. 27, no. 12, pp. 1643–1655, 2007.
- [60] C. Dobler, S. Hagemann, R. L. Wilby, and J. Stötter, "Quantifying different sources of uncertainty in hydrological projections in an Alpine watershed," *Hydrology and Earth System Sciences*, vol. 16, no. 11, pp. 4343–4360, 2012.
- [61] T. Hashino, A. A. Bradley, and S. S. Schwartz, "Evaluation of bias-correction methods for ensemble streamflow volume forecasts," *Hydrology and Earth System Sciences*, vol. 11, no. 2, pp. 939–950, 2007.
- [62] C. Piani, G. P. Weedon, M. Best et al., "Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models," *Journal of Hydrology*, vol. 395, no. 3-4, pp. 199–215, 2010.
- [63] M. J. Themeßl, A. Gobiet, and G. Heinrich, "Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal," *Climatic Change*, vol. 112, no. 2, pp. 449–468, 2012.
- [64] M. Jakob Themeßl, A. Gobiet, and A. Leuprecht, "Empirical-statistical downscaling and error correction of daily precipitation from regional climate models," *International Journal of Climatology*, vol. 31, no. 10, pp. 1530–1544, 2011.
- [65] A. W. Wood, L. R. Leung, V. Sridhar et al., "Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs," *Climatic Change*, vol. 62, no. 1-3, pp. 189–216, 2004.
- [66] R. Huth, "Statistical downscaling of daily temperature in central Europe," *Journal of Climate*, vol. 15, no. 13, pp. 1731–1742, 2002.
- [67] J. Friedman, T. Hastie, H. Höfling, and R. Tibshirani, "Pathwise coordinate optimization," *The Annals of Applied Statistics*, vol. 1, no. 2, pp. 302–332, 2007.
- [68] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society B: Methodological*, vol. 58, no. 1, pp. 267–288, 1996.
- [69] W. J. Fu, "Penalized regressions: the bridge versus the lasso," *Journal of Computational and Graphical Statistics*, vol. 7, no. 3, pp. 397–416, 1998.
- [70] A. M. Ebtehaj, E. Foufoula-Georgiou, and G. Lerman, "Sparse regularization for precipitation downscaling," *Journal of Geophysical Research D: Atmospheres*, vol. 117, no. D8, Article ID D08107, 2012.

